MOLECULAR BIOLOGY OF THE CELL

THIRD EDITION

Bruce Alberts • Dennis Bray
Julian Lewis • Martin Raff • Keith Roberts
James D. Watson

Garland Publishing, Inc.
New York & London

Text Editor: Miranda Robertson
Managing Editor: Ruth Adams
Illustrator: Nigel Orme
Molecular Model Drawings: Kate Hesketh-Moore
Director of Electronic Publishing: John M-Roblin
Computer Specialist: Chuck Bartelt
Disk Preparation: Carol Winter
Copy Editor: Shirley M. Cobert
Production Editor: Douglas Goertzen
Production Coordinator: Perry Bessas
Indexer: Maija Hinkle

Bruce Alberts received his Ph.D. from Harvard University and is currently President of the National Academy of Sciences and Professor of Biochemistry and Biophysics at the University of California, San Francisco. *Dennis Bray* received his Ph.D. from the Massachusetts Institute of Technology and is currently a Medical Research Council Fellow in the Department of Zoology, University of Cambridge. *Julian Lewis* received his D.Phil. from the University of Oxford and is currently a Senior Scientist in the Imperial Cancer Research Fund Developmental Biology Unit, University of Oxford. *Martin Raff* received his M.D. from McGill University and is currently a Professor in the MRC Laboratory for Molecular Cell Biology and the Biology Department, University College London. *Keith Roberts* received his Ph.D. from the University of Cambridge and is currently Head of the Department of Cell Biology, the John Innes Institute, Norwich. *James D. Watson* received his Ph.D. from Indiana University and is currently Director of the Cold Spring Harbor Laboratory. He is the author of *Molecular Biology of the Gene* and, with Francis Crick and Maurice Wilkins, won the Nobel Prize in Medicine and Physiology in 1962.

Library of Congress Cataloging-in-Publication Data
Molecular biology of the cell / Bruce Alberts . . . [et al.].—3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-8153-1619-4 (hard cover).—ISBN 0-8153-1620-8 (pbk.)
1. Cytology. 2. Molecular biology. I. Alberts, Bruce.
[DNLM: 1. Cells. 2. Molecular Biology. QH 581.2 M718 1994]
QH581.2.M64 1994
574.87—dc20
DNLM/DLC
for Library of Congress
93-45907
CIP

Published by Garland Publishing, Inc.
717 Fifth Avenue, New York, NY 10022

Printed in the United States of America

15 14 13 12 10 9 8 7 6 5 4

Front cover: The photograph shows a rat nerve cell in culture. It is labeled (*yellow*) with a fluorescent antibody that stains its cell body and dendritic processes. Nerve terminals (*green*) from other neurons (not visible), which have made synapses on the cell, are labeled with a different antibody. (Courtesy of Olaf Mundigl and Pietro de Camilli.)

Dedication page: Gavin Borden, late president of Garland Publishing, weathered in during his mid-1980s climb near Mount McKinley with MBoC author Bruce Alberts and famous mountaineer guide Mugs Stump (1940–1992).

Back cover: The authors, in alphabetical order, crossing Abbey Road in London on their way to lunch. Much of this third edition was written in a house just around the corner. (Photograph by Richard Olivier.)

acts. If these minor cell proteins differ among cells to the same extent as the abundant proteins, as is commonly assumed, only a small number of differences (perhaps several hundred) suffice to create very large differences cell morphology and behavior.

Cell Can Change the Expression of Its Genes Response to External Signals [3]

st of the specialized cells in a multicellular organism are capable of altering ir patterns of gene expression in response to extracellular cues. If a liver cell exposed to a glucocorticoid hormone, for example, the production of several ecific proteins is dramatically increased. Glucocorticoids are released during riods of starvation or intense exercise and signal the liver to increase the oduction of glucose from amino acids and other small molecules; the set of oteins whose production is induced includes enzymes such as tyrosine aminonsferase, which helps to convert tyrosine to glucose. When the hormone is no nger present, the production of these proteins drops to its normal level.

Other cell types respond to glucocorticoids in different ways. In fat cells, for ample, the production of tyrosine aminotransferase is reduced, while some ther cell types do not respond to glucocorticoids at all. These examples illustrate general feature of cell specialization—different cell types often respond in different ways to the same extracellular signal. Underlying this specialization are ntures that do not change, which give each cell type its permanently distinctive character. These features reflect the persistent expression of different sets of nes.

ene Expression Can Be Regulated at Many of the Steps n the Pathway from DNA to RNA to Protein [4]

f differences between the various cell types of an organism depend on the particular genes that the cells express, at what level is the control of gene expression exercised? There are many steps in the pathway leading from DNA to protein, and all of them can in principle be regulated. Thus a cell can control the proteins it makes by (1) controlling when and how often a given gene is transcribed (**transcriptional control**), (2) controlling how the primary RNA transcript is spliced or otherwise processed (**RNA processing control**), (3) selecting which completed mRNAs in the cell nucleus are exported to the cytoplasm (**RNA transport control**), (4) selecting which mRNAs in the cytoplasm are translated by ribosomes (**translational control**), (5) selectively destabilizing certain mRNA molecules in the cytoplasm (**mRNA degradation control**), or (6) selectively activating, inactivating, or compartmentalizing specific protein molecules after they have been made (**protein activity control**) (Figure 9–2).

For most genes transcriptional controls are paramount. This makes sense because, of all the possible control points illustrated in Figure 9–2, only transcriptional control ensures that no superfluous intermediates are synthesized. In the

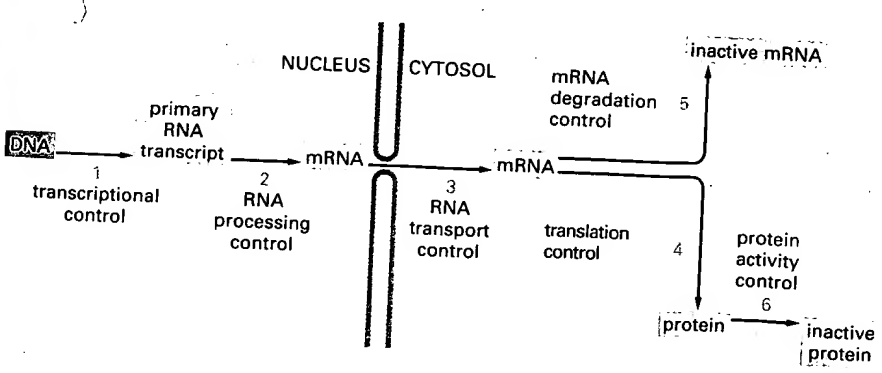

Figure 9–2 **Six steps at which eucaryote gene expression can be controlled.** Only controls that operate at steps 1 through 5 are discussed in this chapter. The regulation of protein activity (step 6) is discussed in Chapter 5; this includes reversible activation or inactivation by protein phosphorylation as well as irreversible inactivation by proteolytic degradation.

following sections we discuss the DNA and protein components that regulate the initiation of gene transcription. We return at the end of the chapter to the other ways of regulating gene expression.

Summary

The genome of a cell contains in its DNA sequence the information to make many thousands of different protein and RNA molecules. A cell typically expresses only a fraction of its genes, and the different types of cells in multicellular organisms arise because different sets of genes are expressed. Moreover, cells can change the pattern of genes they express in response to changes in their environment, such as signals from other cells. Although all of the steps involved in expressing a gene can in principle be regulated, for most genes the initiation of RNA transcription is the most important point of control.

DNA-binding Motifs in Gene Regulatory Proteins [5]

How does a cell determine which of its thousands of genes to transcribe? As discussed in Chapter 8, the transcription of each gene is controlled by a regulatory region of DNA near the site where transcription begins. Some regulatory regions are simple and act as switches that are thrown by a single signal. Other regulatory regions are complex and act as tiny microprocessors, responding to a variety of signals that they interpret and integrate to switch the neighboring gene on or off. Whether complex or simple, these switching devices consist of two fundamental types of components: (1) short stretches of DNA of defined sequence and (2) *gene regulatory proteins* that recognize and bind to them.

We begin our discussion of gene regulatory proteins by describing how these proteins were discovered.

Gene Regulatory Proteins Were Discovered Using Bacterial Genetics [6]

Genetic analyses in bacteria carried out in the 1950s provided the first evidence of the existence of **gene regulatory proteins** that turn specific sets of genes on or off. One of these regulators, the *lambda repressor,* is encoded by a bacterial virus, *bacteriophage lambda.* The repressor shuts off the viral genes that code for the protein components of new virus particles and thereby enables the viral genome to remain a silent passenger in the bacterial chromosome, multiplying with the bacterium when conditions are favorable for bacterial growth (see Figure 6–80). The lambda repressor was among the first gene regulatory proteins to be characterized, and it remains one of the best understood, as we discuss later. Other bacterial regulators respond to nutritional conditions by shutting off genes encoding specific sets of metabolic enzymes when they are not needed. The *lac repressor,* for example, the first of these bacterial proteins to be recognized, turns off the production of the proteins responsible for lactose metabolism when this sugar is absent from the medium.

The first step toward understanding gene regulation was the isolation of mutant strains of bacteria and bacteriophage lambda that were unable to shut off specific sets of genes. It was proposed at the time, and later proved, that most of these mutants were deficient in proteins acting as specific repressors for these sets of genes. Because these proteins, like most gene regulatory proteins, are present in small quantities, it was difficult and time-consuming to isolate them. They were eventually purified by fractionating cell extracts on a series of standard chromatography columns (see pp. 166–169). Once isolated, the proteins were shown to bind to specific DNA sequences close to the genes that they

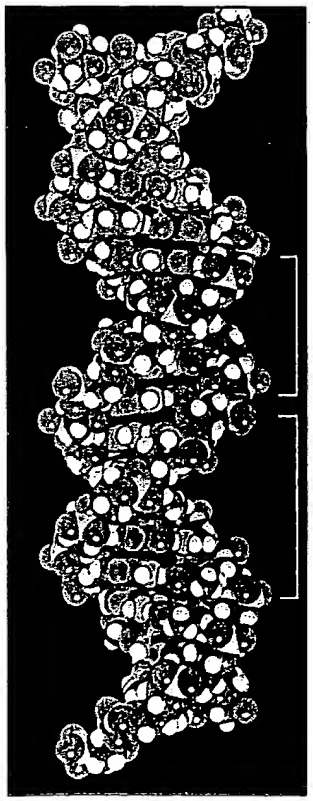

Figure 9–3 **Double-helical structure of DNA.** The major and minor grooves on the outside of the double helix are indicated. The atoms are colored as follows: carbon, *dark blue;* nitrogen, *light blue;* hydrogen, *white;* oxygen, *red;* phosphorus, *yellow.*

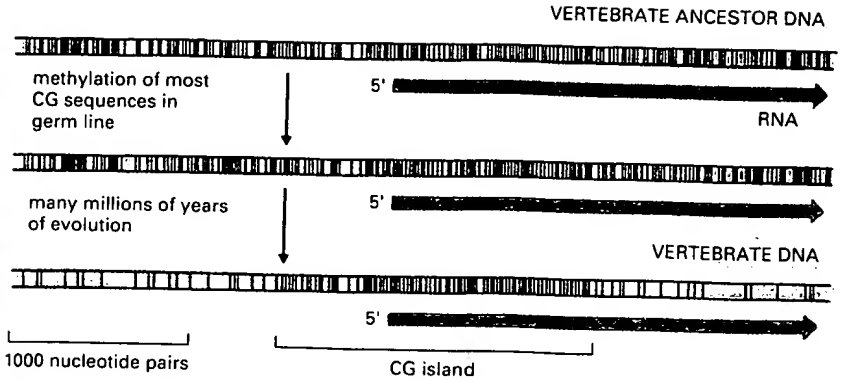

Figure 9–71 A mechanism to explain both the marked deficiency of CG sequences and the presence of CG islands in vertebrate genomes. A *black line* marks the location of an unmethylated CG dinucleotide in the DNA sequence, while a *red line* marks the location of a methylated CG dinucleotide.

Summary

The many types of cells in animals and plants are created largely through mechanisms that cause different genes to be transcribed in different cells. Since many specialized animal cells can maintain their unique character when grown in culture, the gene regulatory mechanisms involved in creating them must be stable once established and heritable when the cell divides, endowing the cell with a memory of its developmental history. Procaryotes and yeasts provide unusually accessible model systems in which to study gene regulatory mechanisms, some of which may be relevant to the creation of specialized cell types in higher eucaryotes. One such mechanism involves a competitive interaction between two (or more) gene regulatory proteins, each of which inhibits the synthesis of the other; this can create a flip-flop switch that switches a cell between two alternative patterns of gene expression. Direct or indirect positive feedback loops, which enable gene regulatory proteins to perpetuate their own synthesis, provide a general mechanism for cell memory.

In eucaryotes gene transcription is generally controlled by combinations of gene regulatory proteins. It is thought that each type of cell in a higher eucaryotic organism contains a specific combination of gene regulatory proteins that ensures the expression of only those genes appropriate to that type of cell. A given gene regulatory protein may be expressed in a variety of circumstances and typically is involved in the regulation of many genes.

In addition to diffusible gene regulatory proteins, inherited states of chromatin condensation are also utilized by eucaryotic cells to regulate gene expression. In vertebrates DNA methylation also plays a part, mainly as a device to reinforce decisions about gene expression that are made initially by other mechanisms.

Posttranscriptional Controls

Although controls on the initiation of gene transcription are the predominant form of regulation for most genes, other controls can act later in the pathway from RNA to protein to modulate the amount of gene product that is made. Although these **posttranscriptional controls,** which operate after RNA polymerase has bound to the gene's promoter and begun RNA synthesis, are less common than *transcriptional control,* for many genes they are crucial. It seems that every step in gene expression that could be controlled in principle is likely to be regulated under some circumstances for some genes.

We consider the varieties of posttranscriptional regulation in temporal order, according to the sequence of events that might be experienced by an RNA molecule after its transcription has begun (Figure 9–72).

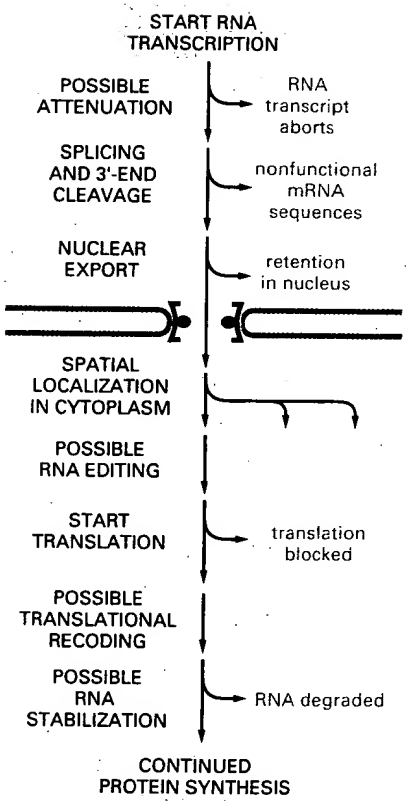

Figure 9–72 Possible posttranscriptional controls in gene expression. Only a few of these controls are likely to be used for any one gene.